1980-06

An analysis of accidents involving towboat-barge combination on selected inland waterways of the United States.

Gamble, William John

http://hdl.handle.net/10945/17512

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

AN ANALYSIS OF ACCIDENTS INVOLVING
TOWBOAT-BARGE COMBINATIONS ON SELECTED
INLAND WATERWAYS OF THE UNITED STATES

by

William John Gamble

June, 1980

Thesis Advisor: D. E. Neil

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Analysis of Accidents Involving Towboat-Barge Combinations on Selected Inland Waterways of the United States		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis June 1980
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William John Gamble		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE June 1980
		13. NUMBER OF PAGES 89
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

ACCIDENTS; TOWBOAT; BARGE; TOWBOAT-BARGE COMBINATIONS; WATERWAYS; INLAND WATERWAYS; UNITED STATES

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study uses a statistical analysis approach on a computerized data base to analyze accidents involving towboat-barge combinations on the inland waterways of the United States. The main areas explored are the factors affecting the severity and the frequency of accidents. In addition, multiple regression models are used to predict the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
(Page 1) S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

BLOCK 20, ABSTRACT (Cot'd.)

severity of towboat accidents from a set of independent accident variables. Conclusions and recommendations are given on towboat accidents and on marine accident data collection and analysis.

An Analysis of Accidents Involving Towboat-Barge Combinations on Selected Inland Waterways of the United States

by

William John Gamble
Lieutenant, United States Coast Guard
B.S., United States Coast Guard Academy, 1971

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June, 1980

ABSTRACT

This study uses a statistical analysis approach on a computerized data base to analyze accidents involving towboat-barge combinations on the inland waterways of the United States. The main areas explored are the factors affecting the severity and the frequency of accidents. In addition, multiple regression models are used to predict the severity of towboat accidents from a set of independent accident variables. Conclusions and recommendations are given on towboat accidents and on marine accident data collection and analysis.

TABLE OF CONTENTS

LIST OF TABLES

table *page*

LIST OF FIGURES

I. INTRODUCTION

The aim of this thesis is to use simple statistical tests and models to study accidents involving towboat-barge combinations operating on the inland waters of the United States. The thesis begins with a description of the towboat-barge transportation system and the problem of accidents in that system. Next, data available on towboats and towboat accidents is presented. Following this is the application of various statistical tests and models to the data. Finally, conclusions about the effort are drawn.

Before discussing the problem of accidents in the towboat-barge transportation system, it is appropriate to discuss marine accidents in general. Marine accidents and casualties pose a serious problem. Coast Guard commercial marine accident statistics show that for fiscal year 1978, rammings, groundings, collisions, and other vessel casualties amounted to 4,268 incidents with 7,118 vessels involved and 179 deaths. These figures do not take into account recreational-boating accidents, which, in calendar year 1978, resulted in 6,529 incidents with 8,576 boats involved and 1,321 deaths.

Monetary losses from commercial vessel casualties are also high. For fiscal year 1978, two hundred million dollars were lost due to vessel and cargo damage. It is very likely that this figure does not reflect the true cost of marine casualties. The true cost includes more than just vessel and cargo damage. Besides the cost of vessel damage, there are the additional costs of mooring an idle vessel, paying an idle crew, and the penalty cost of lost revenue. Besides the cost of cargo damage, there may be large clean-up costs, for example, when petroleum products, chemical products, or other hazardous liquid cargoes spill into the water.

The Federal agency primarily responsible for controlling the problem of accidents in the marine community is the U. S. Coast Guard. In fact, the main goal of the Coast Guard is to minimize loss of life, personal injury, and property damage on, over, and under the high seas and waters subject to U. S. jurisdiction. This thesis was undertaken in the spirit of this goal, and in the hope of contributing to a better understanding of accidents involving a particular segment of the marine community, i.e., towboats and their barge combinations on the inland waters of the U. S.

II. THE TOWBOAT-BARGE TRANSPORTATION SYSTEM

A. THE TOWBOATS

Since the towboat is designed for the express purpose of pushing barges on the inland waterways, it has a rather unconventional shape (see Figure 1). The bow is flat, with two uprights, called towing knees, protruding from it. Barges are secured against these knees for pushing. In the forepart of the vessel, a few feet back from the bow, the superstructure rises directly to the highest point on the vessel, the pilot house. From the pilot house the operator controls the vessel and maintains alignment of the barges being pushed. Behind the pilot house, the superstructure drops rapidly to give the operator adequate stern view. Located in the middle of the vessel are the stacks which emit waste gases from the engine. The vessel is finished off with a box-shaped stern. Some common sizes of towboats are given in Table 1, below.

Table 1 Sizes of Towboats

size	*length, feet*	*width, feet*	*draft, feet*	*horsepower*	*barges*
Small	117	30	7.6	1,000-2,000	8
Medium	124	34	8.0	2,000-4,000	16
Large	160	40	8.6	4,000-6,000	24

Figure 1 Towboat Without a Barge Combination
(Photograph courtesy of American Waterways Operators, Inc.)

A distinction should be made at this point between a towboat and a tugboat. Although both of these vessels come under the general heading of a towing vessel, the towboat pushes barges while the tugboat pulls or tows them. In the push method, the barges are lashed together to form a single unit, and then secured to the towing knees of the towboat. In the pull method, the barges are secured to a hawser behind the tugboat and then pulled to their destination. The push method is more effective, and provides greater control of the load in calm water. This method is therefore more suitable for inland waters which are naturally calm or calmed by a series of locks and dams. On the other hand, the pull method is reserved for open-ocean towing in which the water is too rough to keep the barges lashed together.

The number of barges that a towboat may push varies according to the size of the towboat, environmental conditions, and limitations on the inland waterway. Under normal conditions, the towboat industry has a rule of thumb for the maximum number of barges that can be pushed. This rule is 250 horsepower for each barge in the towboat-barge combination. The final column in Table 1 uses this rule to compute the maximum number of barges these vessels would normally push.

There are towboats both smaller and larger than the common sizes given above. In fact, the smallest towboats are about 36 feet in length and produce around 100 horsepower, while the largest towboats exceed 170 feet in length and produce over 9,000 horsepower. It is also interesting to note that towboats are increasing in horsepower. According to American Waterways Operators, Inc., in 1962 the average was 672 horsepower, while in 1972 average towboat horsepower was 1,006. [Ref. 1]

The equipment on board towboats is similar to that found on most vessels, and includes such items as radar, radiotelephone, depth finder, automatic pilot, and search lights. Occasionally, the equipment is modified for towboat use. For example, the depth finder on some towboats can be operated by a transceiver suspended in the water from the lead barge in the towboat-barge combination. This modification gives the operator a better indication of upcoming changes in water depth. One piece of special equipment found on certain towboats is a swing meter to monitor the alignment of the towboat-barge combination.

Besides the above-water differences between towboats and other vessels, there are underwater differences. Unlike most vessels, a towboat has a flat bottom and from two to four propellers. Fore and aft of the propellers is a series of controllable rudders. This makes maneuvering a towboat different from maneuvering other vessels. In fact, on towboats, there is no wheel or helm controlling the rudders; instead, a series of handles or levers is used (see Figure 2). The steering process on the towboat consists of adjusting the levers to get the proper angle on the rudders, and adjusting the engine rpm to control the thrust from the screws.

The speed attained by a towboat depends on the environmental conditions and the number of barges being pushed. A rough average speed for a towboat-barge combination is about six knots, with a maximum speed of perhaps fifteen knots. At these speeds it takes a considerable amount of time for a towboat-barge combination to make a trip. Reference 1 gives typical transit times for a towboat-barge combination over various inland waterways. A few of these times are reproduced in Table 2, below.

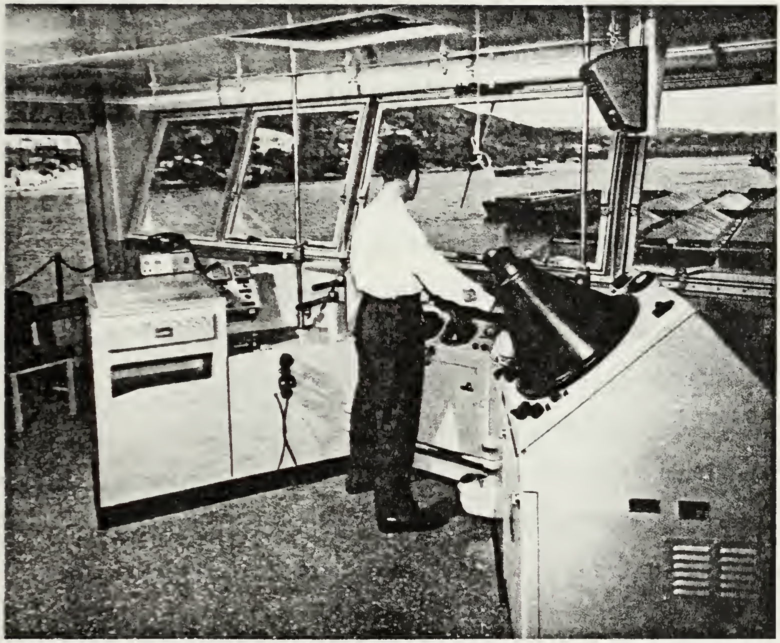

Figure 2 Pilot in the Pilot House Maneuvering a Towboat
(Photograph courtesy of American Waterways Operators, Inc.)

Table 2 Transit Times for an Average Towboat-Barge Combination

from	*to*	*distance, miles*	*upstream time*	*downstream time*
Chicago	New Orleans	1,418	11 days 8 hrs.	6 days 7 hrs.
Cincinnati	Houston	1,785	13 days 15 hrs.	8 days 8 hrs.
Pittsburgh	Brownsville	2,542	18 days 21 hrs.	12 days 8 hrs.

Just as important as making headway is being able to stop. Like making headway, stopping depends on the environmental conditions and the load of the towboat. According to American Waterways Operators, Inc., under perfect environmental conditions, and in calm water, an average towboat-barge combination can be stopped in one and one-half towboat-barge lengths. [Ref. 1]

The U. S. Coast Guard keeps records on commercial vessels registered in the United States. These records give the number of vessels working in the towing industry. The towing industry includes both towboats and tugboats on all U. S. waterways. As can be seen in Table 3, below, the number of registered towing vessels has been steadily growing in recent years.

Table 3 Number of Registered Towing Vessels

Year	*Number*	*Year*	*Number*
1971	6,039	1975	6,549
1972	6,057	1976	6,705
1973	6,149	1977	6,813
1974	6,308		

B. THE BARGES

A barge is a special vessel designed for the express purpose of carrying cargo. It is of welded-steel construction, and is boxlike in shape to permit carrying the maximum amount of cargo. Normally, the bow is raked or sloped to permit ease of movement through the water. The stern, on the other hand, is boxed, or square in shape, for ease of pushing (see Figures 3 and 4).

Depending on the purpose of the barge, the overall shape may be modified. In large tows, where many barges are lashed together, the middle barges may be boxed on both ends. The placing of square ends together in the middle of the tow makes a smooth underwater body that reduces water resistance. Smaller tows, on the other hand, do not normally have these special barges. In small tows, boxed ends are frequently placed against raked ends, increasing the water resistance of the tow.

There are three common types of barges found on the inland waterways: the hopper barge, the deck barge, and the tank barge. Hopper barges range from 175 to 290 feet in length and from 26 to 50 feet in width. The draft of these barges when loaded is about nine feet. A hopper barge is basically a box without a top, the only difference being that there is an inner and an outer skin. The inner skin forms the hopper or hold, and the outer skin forms the exterior of the barge. Between the two skins are voids, or pockets of air, which can protect the barge from flooding and sinking in the event of a collision. The largest voids are found at the bow and at the stern. The hopper barge can carry a bulk cargo, such as coal, or a non-bulk cargo, such as a finished good. The cargo capacity of such barges ranged from 1,000 to 3,000 tons. Sometimes a cover or top is placed over the barge to protect the cargo, forming a so-called dry-cargo barge.

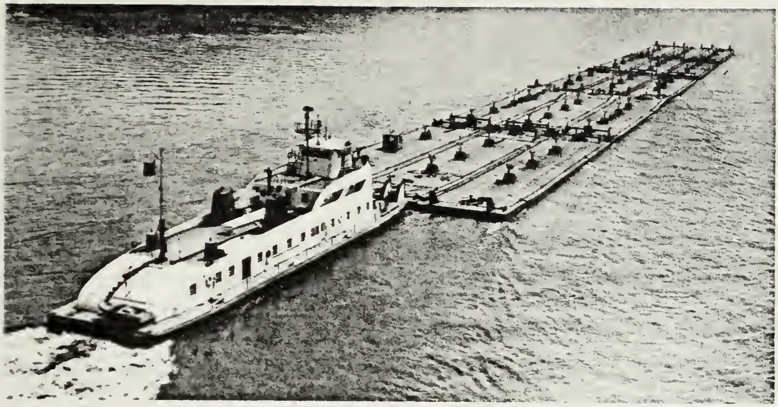

Figure 3   Tank Barges Being Pushed by a Towboat
(Photograph courtesy of American Waterways Operators, Inc.)

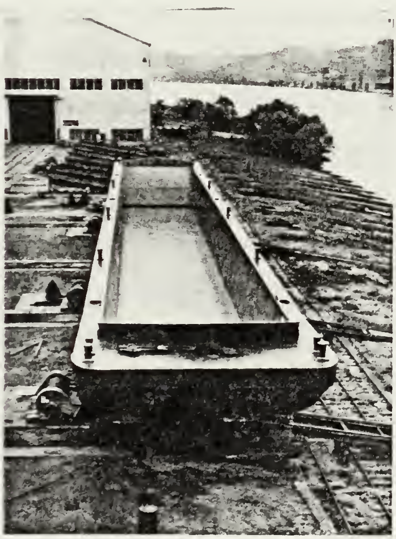

Figure 4   Hopper Barge in a Shipyard
(Photograph courtesy of American Waterways Operators, Inc.)

Deck barges range from 110 to 195 feet in length, from 26 to 35 feet in width, and have a draft of approximately eight feet. They are designed with a heavy well-supported deck that is either used to transport heavy equipment or used as a working platform. The carrying capacity of these barges ranges from 350 to 1200 tons.

Tank barges are used to transport bulk liquids. They generally range from 175 to 290 feet in length, from 26 to 50 feet in width, and have a nine-foot draft. Their cargo capacity ranges from 300,000 to 900,000 gallons. Three common types of tank barge are the single-skinned barge, the double-skinned barge, and the barge with independent cylindrical tanks. The type of liquid to be transported dictates the type of barge used. Liquids that pose no hazard to the environment can be transported in a single-skinned barge. This barge would normally have both a bow and a stern void, but no voids along the sides; thus, a collision which penetrated the side of such a barge would release the cargo. Double-skinned barges are used to transport more hazardous cargoes: voids completely surround the cargo, thus protecting it during a collision. Volatile liquids that need to be transported under reduced temperatures or under high pressures are placed in barges with independent cylindrical tanks.

Barges are put together in tows for ease of pushing by the towboat. Wire rope and line looped around deck bitts are used to join the barges together. While in the tow and while being handled, barges frequently sustain minor damage. Weld fractures and small punctures in the hull are the most common result.

C. THE PERSONNEL

The most important person on the towboat is the operator or pilot. He is located in the pilot house and it is his responsibility to control the movement of the towboat-barge combination. His job is difficult, and to do it properly requires a combination of natural ability, training, and experience.

It is interesting to note the differences between the job of a towboat operator and the job of a watch officer on a deep-sea vessel. On the towboat the operator is usually the only person in the pilot house, while on a deep-sea vessel the watch officer may have several assistants on the bridge. The watch rotation on towboats is also unlike that on deep-sea vessels. An operator on a towboat follows a schedule of six hours on duty and six hours off, while an officer on a deep-sea vessel has a schedule of four hours on duty and eight hours off.

The tasks a towboat operator is required to perform are quite varied and depend on the maneuvering situation. For example, in maneuvering into a berth, he would be doing such tasks as monitoring the position of the towboat-barge combination in relation to the berth, keeping track of the effects of wind and current, adjusting the towboat's engine and rudder controls, watching gauges and instruments, sounding the proper signals, and, if necessary, monitoring or talking on the radiotelephone. Although intense concentration is required for these tasks, the concentration is only necessary for short periods of time (e.g., the time required to maneuver the vessel into a berth). In a situation such as maintaining course and speed over a stretch of navigable channel, the concentration is less intense but more prolonged. In this situation, the operator would still

be doing such tasks as adjusting the engine and rudder controls, keeping track of the effects of wind and current, watching the readings on gauges and instruments, and either monitoring or using the radiotelephone. There are a few different tasks required in navigating on open channels; these would include studying the intended track, evaluating approaching traffic, and evaluating upcoming obstructions, both natural (such as bends in the channel) and man-made (such as bridges).

In 1972 Public Law 92-339 was passed. This law set a maximum of 12 hours of work during any 24-hour period on board towboats, and required operators of towboats to pass an examination for a license. The regulations to implement this law were drafted by the Coast Guard and became effective in September of 1973. Personnel already serving as operators of towboats were exempted from the test requirement, if they had sufficient documented past experience. The present formal test for a towboat license covers such items as using navigational instruments, using charts for navigation, and understanding the rules of the road.

Another piece of legislation which affected towboat operators was the so-called "Bridge-to-Bridge Radiotelephone Act" (Public Law 92-63). It became effective on 1 January 1973, and required every towing vessel of 26 feet or more in length to carry a radiotelephone for the exchange of navigational information. This act was important to inland waterway operators because it promoted the exchange of navigational information between vessels before they met or crossed.

The living and working conditions of towboat personnel vary according to the towboat, the towboat company, and the route or trip taken. Normally, crew members on towboats work an 84-hour week--six hours on duty and six

hours off duty throughout the voyage. Depending on the company, the employee may receive from one-third of a day to a full day off for each day worked. As on most commercial vessels, living conditions on board towboats are difficult. The towboat has a large powerplant in comparison to its size, and this both cramps the living quarters and increases the ambient noise levels. However, newer towboats have been designed and constructed with more concern for crew living conditions.

D. THE INLAND WATERWAYS

The inland waterways of the United States have about 25,000 miles of usable navigational channels, exclusive of the Great Lakes. About 15,000 miles have an operating depth of nine feet or more, with the remainder being shallower. Figure 5 shows the location of these channels in the eastern half of the United States.

Most of the channels are rivers that must be kept in navigable condition. The two government agencies primarily responsible for this are the U. S. Army Corps of Engineers and the U. S. Coast Guard. The Coast Guard places and maintains various aids to navigation along the channels. Typical aids are buoys and lights that both mark the safe area of the channel and mark obstructions in the channel. The Army Corps of Engineers is responsible for maintaining the navigable characteristics of the channels. This includes maintaining the channels at their operating depths, removing obstructions in the channel, and in general improving the channel where necessary.

An important part of the navigation of inland waterways is the lock and dam system. Locks and dams help maintain a constant depth, and permit vessels to safely transit changes in the elevation of the channel.

Figure 5 Inland Waterways in the Eastern Half of the United States
(Chartlet courtesy of American Waterways Operators, Inc.)

A lock is essentially a chamber in which the water level can be adjusted. A towboat-barge combination is placed in a lock and the water level is either raised or lowered to meet the change in the river's height. The size of a lock may therefore be a limiting factor in the size of the combinations that can transit a channel. In Table 4, below, common sizes of locks are given, together with the maximum sizes of towboat-barge combinations that can fit into each.

Table 4 Sizes of Common Locks

lock width	*lock length*	*barges*	*towboats*
110 ft.	1,200 ft.	20	1
110 ft.	600 ft.	10	1

Towboat-barge combinations that are too big to fit through a lock must be broken up into smaller sections. Each section is then passed through separately, and the towboat-barge combination is put together on the other side. A towboat-barge combination that does not have to be broken up can be passed through a lock in about half an hour. According to Howe [Ref. 3], a towboat-barge combination that has to be broken into two sections takes about an hour and a half to pass through a lock.

The data in this study was collected on four of the major inland waterways. These waterways are the Mississippi, Illinois, Ohio, and Gulf Intracoastal. Table 5, below, gives some information on each of these waterways.

Table 5 Waterways on which Accident Data is Available

name	*length, miles*	*number of locks*	*principal commodities transported*
Mississippi	2,360	30	petroleum, grain, aluminum
Illinois	354	7	coal, petroleum, grain, iron
Ohio	981	43	coal, petroleum, gravel, chemicals
Gulf Intracoastal	1,113	9	petroleum, chemicals, iron, sea shells

Another feature of the inland waterways that should be noted is that there are various plants and terminals located along most of them. Plants include such structures as oil refineries, cement plants, iron and steel mills, power plants, aluminum plants, glass plants, fertilizer plants, etc. Terminals are of two basic types, bulk-loading and non-bulk-loading. Examples of bulk-loading terminals would include coal, grain, and petroleum terminals. Examples of non-bulk-loading terminals would be docks to off-load special cargoes such as industrial equipment or steel pipe.

E. THE ACCIDENT PROBLEM

According to American Waterways Operators, Inc. [Ref. 1], approximately 1,800 companies are engaged in commercial transportation of commodities on the inland waterways. About 80,000 persons are employed on board the inland fleet, and about an equal number of persons are employed in shore-based support work. Fleet personnel include operators, engineers, deck hands, cooks, etc. Shore-based personnel include office workers, service workers, and shipbuilding and repair workers.

Some information on the safety record of the towing industry is available from Coast Guard marine casualty records. The towing industry includes both towboats and tugboats operating on inland waterways and offshore. The information can be divided into two basic categories: accidents considered vessel casualties and accidents considered non-vessel casualties. A vessel casualty would be an accident that affects the seaworthiness of the vessel. Examples include rammings, groundings, and collisions. A non-vessel casualty would be an injury or death on board the vessel in which the vessel's seaworthiness was not affected. Examples include an injury or death caused by an electric shock, a heart attack, or a slip and subsequent fall.

Only the more important marine accidents are reported to the Coast Guard. Reportable marine accidents are defined by Federal regulations as accidents in which one of more of the following have occurred:

(a) actual physical damage to property in excess of $1500;

(b) material damage to the vessel affecting its seaworthiness;

(c) stranding or grounding of the vessel;

(d) loss of life; or

(e) injury causing any person to remain incapacitated for a period in excess of 72 hours.

Coast Guard statistics give information on vessel casualties for three types of vessels associated with the towing industry: inspected cargo barges, inspected tank barges, and towing vessels. For the past five years the average number of deaths and injuries due to vessel casualties from these three types of vessels has been quite low. The average number of deaths per year is 20, and the average number of injuries is 21. For the past four years, the average number of deaths and injuries due to non-vessel casualties

from towing vessels only is also available. The averages are 43 deaths per year and 89 injuries per year.

Coast Guard statistics give the number of towing vessels involved in reportable casualties for the past five years. The definition of a reportable casualty was given above. Using a yearly average, the approximate number of towboats and tugboats involved in reportable casualties per year for the past five years is 1,480. For approximately the same period the yearly average of the number of towing vessels registered with the Coast Guard was 6,505. Discounting the possibility that some towing vessels are involved in more than one accident per year, it can be said that roughly 20% of the towboats and tugboats have some type of ramming, grounding, or other vessel casualty during a year.

Although information on the dollar value of damage in towboat accidents is not directly available from Coast Guard statistical publications, a mean dollar-value figure per accident for the period 1971-1976 was computed from the data base on towboat accidents used for this study. The figure is $53,291, using 1978 as a base year. This figure is the Coast Guard investigating officer's estimate of the damage, and includes the dollar value of damage to the towboat, barges, cargo, and other property such as docks or bridge pilings. Although quite high, this figure does not give the real cost of a towboat accident. The real cost includes not only damage to vessel, cargo, and property, but more importantly the cost of lost revenue while the damaged vessel is being repaired, and the large costs that may accrue when petroleum or chemical products spill into the water and clean-up operations are necessary. In addition, there are costs that are difficult to measure, such as the adverse publicity that surrounds

an accident. In actual fact, it is very likely that the total seen and unseen costs amount to a figure much higher than $53,291.

As described above, both the large number and the severity of accidents create financial hardships for the towboat-barge transportation system. Coast Guard investigating officers and industry representatives can affirm that there is no simple solution. The only hope is that through continued study of the problem, methods of solving it will evolve.

After many computer runs and repeated application of statistical techniques, it was realized that statistical analysis also cannot find the best approach to the problem. As a second-best alternative, the statistical techniques applied were designed to increase understanding of the problem. In particular, it is hoped that this approach will provide a different perspective from which to view the problem of towboat accidents.

III. DESCRIPTION OF THE DATA

The data used in this thesis came from two sources: the U. S. Coast Guard and the U. S. Army Corps of Engineers. This chapter will describe the data obtained from these sources and relate how that data was prepared for analysis.

The Army Corps of Engineers collects data on the number of towboats and barges making trips on the various channels of the inland waterway system. The data is compiled in an Army Corps of Engineers publication entitled *Waterborne Commerce of the United States* [Ref. 9]. The data is split into upbound and downbound trips. A figure representing the total number of trips was computed by adding the upbound and downbound trips together. The data is displayed in Table 6, below. Unfortunately, data for all four of the navigational channels used in this study could not be extracted from Reference 9. The data for the Gulf Intracoastal Waterway was not available.

Table 6 Barges/Towboats Making Trips on the Inland Waterways

Year	*Mississippi Waterway*	*Illinois Waterway*	*Ohio Waterway*
1972	249,619/88,360	47,166/8,173	227,102/67,204
1973	228,956/79,869	46,743/8,286	232,548/67,478
1974	248,004/86,080	46,316/8,496	223,331/67,331
1975	243,081/82,066	49,683/8,597	215,008/64,378
1976	259,614/79,723	50,751/7,383	229,851/65,097
1977	263,452/82,302	48,294/7,714	233,561/65,533

The data used from the U. S. Coast Guard includes both statistical publications and computer tapes. The statistical publications used were *Statistics of Casualties* [Ref. 14], *Marine Safety Statistical Review 1979* [Ref. 13], and *Boating Statistics 1978* [Ref. 10]. In addition to the statistical publications, the number of documented vessels engaged in the trade of towing from 1970 to 1977 was obtained from the Coast Guard's Office of Merchant Marine Safety.

Two computer tapes were obtained from the Coast Guard. The first tape is the Coast Guard's marine casualty data base. On this tape is information on marine casualties from 1962 to 1978. This tape was coded from public casualty reports submitted to the Coast Guard. The second tape consists of towboat accident data collected from various segments of the Mississippi, Ohio, Illinois, and Gulf Intracoastal Waterways. This tape was coded from Coast Guard investigations and other detailed reports on towboat accidents.

A data base upon which to do the analysis was created by combining information from both of these tapes. This combined data base has one record per accident. The data items available on each accident are listed in Table 7, below.

Table 7 List of Data Items

#	Data	#	Data	#	Data
1	Coast Guard official case number	24	Configuration of the combination	45	Time of day
2	Coast Guard official vessel number	25	Number of barges in the combination	46	Visibility
3	Month of casualty	26	Number of loaded barges	47	Wind speed
4	Day of casualty	27	Number of light barges	48	Weather
5	Year of casualty	28	Total cargo tonnage	49	Wind direction
6	Type of casualty	29	Combination length	50	Radar in use?
7	River location of casualty	30	Combination width	51	Radio in use?
8	River milepoint location of casualty	31	Maximum draft	52	Company owning boat
9	Age of operator	32	Channel width at location of accident	53	Number of crew killed
10	Month of operator's birth	33	High/low water at time of casualty	54	Number of passengers killed
11	Day of operator's birth	34	Number of fixed-span bridges within .5 mile	55	Number of longshoremen killed
12	Year of operator's birth	35	Number of movable-span bridges within .5 mile	56	Number of other type persons killed
13	Operator's years of experience	36	Number of locks and dams within .5 mile	57	Number of crew members injured
14	Operator's hours on duty	37	Number of dikes within .5 mile	58	Number of passengers injured
15	Gross tonnage of towboat	38	Number of river bends within .5 mile	59	Number of longshoremen injured
16	Towboat length	39	Number of bars, islands, rocks within .5 mile	60	Number of other type persons injured
17	Towboat horsepower	40	Number of docks within .5 mile	61	Estimated total dollar damage to all vessels involved
18	Number of propellers	41	Number of man-made structures within .5 mile	62	Estimated total dollar damage to all cargo involved
19	Number of flanking rudders	42	Number of canals/navigable rivers within .5 mile	63	Estimated total dollar damage to all property involved
20	Did machinery or equipment failure contribute to casualty?	43	Number of major ports within .5 mile	64	Number of vessels damaged, according to Coast Guard's reporting requirements
21	Towboat draft	44	Maximum span of bridge struck during casualty		
22	Year built				
23	Direction of movement				

Before attempting statistical testing and statistical modelling, it is worthwhile to provide some elementary statistics on the more important data items in the created data base. Elementary statistics on selected data items are given in Table 8, below.

Table 8 Statistics on Selected Data Items

item	*data*	*mean*	*median*	*standard deviation*
9	Age of the operator	41.6 years	--	10.94 years
13	Operator's years of experience	14.09 years	--	9.91 years
14	Operator's hours on duty	2.96 hours	2.96 hours	1.77 hours
17	Horsepower of the towboat	2397.43 hp	--	1781.09 hp
21	Draft of the towboat	7.97 feet	8.48 feet	1.51 feet
25	Number of barges in the towboat-barge combination	6.44	3.97	6.19
26	Number of loaded barges in the combination	4.55	2.68	5.71
30	Width of the towboat-barge combination	76.16 feet	--	39.07 feet
32	Width of the channel at accident location	648.81 feet	--	760.73 feet
44	maximum span of bridge struck during casualty	304.68 feet	--	274.49 feet

Some of the data items used in the analysis were nominal or categorical in nature, and statistics could not be computed on these items. Instead, frequency counts were obtained on some of the more important ones. Table 9, below, lists these data items. Some of them do not have all of the 574 accident records, due to accidents within the data item being missing or miscoded.

Table 9 Frequency Counts on Selected Nominal Scale Data Items

item	*data*	*category*	*freq.*	*pct.*
6	Type of casualty	Head-on collisions	72	12.5
		Bridge rammings	179	31.2
		Lock and dam rammings	111	19.3
		Groundings	91	15.9
		Other	121	21.1
7	River location of the casualty	Lower Mississippi	53	9.2
		Upper Mississippi	96	16.8
		Ohio	120	20.9
		Illinois	89	15.6
		Gulf Intracoastal	215	37.5
46	Visibility at time of accident	.25 mile or less	40	7.5
		.25 mile to .5 mile	16	3.0
		.5 mile to 1 mile	15	2.8
		1 mile to 2 miles	71	13.3
		greater than 2 miles	391	73.4
48	Weather at the time of the casualty	Clear	360	64.5
		Partly cloudy	65	11.6
		Overcast	49	8.8
		Fog	42	7.5
		Rain	24	4.3
		Snow	7	1.3
		Other	11	2.0

In addition to the data items on the combined data base, new items were created for use in the analysis. These new items are arithmetic combinations of the original items. For example, a new item called Item 67 was created by subtracting the width of the towboat-barge combination from the width of the channel. Table 10, below, shows the transformations used and lists some statistics on the new items.

Table 10 Items Created for the Analysis

item	*transformation*	*mean*	*standard deviation*
65	Item 17 ÷ Item 25 (towboat horsepower divided by number of barges)	473.24	324.06
66	Item 17 ÷ Item 26 (towboat horsepower divided by number of loaded barges)	555.86	569.03
67	Item 32 - Item 30 (channel width minus width of the towboat-barge combination)	558.51	726.52
68	Item 44 - Item 30 (maximum span of bridge struck minus towboat-barge combination width)	228.61	255.62

In order to study the severity of towboat accidents, three primary measures of severity were developed and used in the statistical analysis. The first measure was the number of vessels damaged in a towboat casualty; the second was the total dollar damage to all vessels involved in the casualty; and the third was the total dollar damage of the casualty. It should

be noted that the second and third measures of severity are not the exact dollar damage figures, but estimates of those figures by marine investigating officers of the Coast Guard.

The first measure of severity for an accident is Item 64, without any modifications. This item was created simply by counting up the number of records on the first tape mentioned, the Coast Guard's marine casualty data base. Since each record on this tape was one vessel, the total of the records for one casualty gave the number of vessels involved in the casualty that met the Coast Guard's reporting requirements.

The second and third measures of severity came from Items 61 through 63. The second measure was simply Item 61, the total dollar damage to all vessels involved in the casualty. The third measure was created by adding Items 61 through 63 together, giving the total estimated dollar damage figure for the accident.

In order for the second and third measures to severity to be realistic, it was necessary to adjust for inflation. This adjustment was made by converting the dollar damage figure for each year to a base year of 1967. The mean or average of the consumer price index and the producer price index was used for the conversion. These indexes were obtained from the *Statistical Abstract of the United States, 1978* [Ref. 15]. The exact conversion was accomplished by multiplying a conversion factor based on the year the casualty occurred by the dollar damage figures for each accident. The conversion figures are given in Table 11, below.

Table 11
Factors Used to Convert Dollar Damage per Accident to a 1967 Base Year

Year	*Factor*	*Year*	*Factor*
1971	.851	1974	.651
1972	.819	1975	.597
1973	.747	1976	.551

Statistics for the measures of severity are given in Table 12, below. Item 64 is the number of vessels involved in the casualty that met Coast Guard reporting requirements. Item 69 is the total dollar damage figure for the accident. Item 61 is the total dollar damage to all vessels involved in the casualty. Both items are expressed in terms of thousands of dollars, converted to a 1967 base year. Thus the mean of Item 69, 26.62, actually represents $26,619 in 1967 dollars.

Table 12 *Statistics on the Measures of Severity*

Item	*Data*	*Mean*	*Median*	*Standard Deviation*
61	Dollar damage in thousands to vessels	13.42	--	33.68
64	Number of reportable vessel casualties	2.81	2.40	1.48
69	Dollar damage in thousands to vessels, property, and cargo	26.62	5.98	156.65

In any study using statistical analysis of data, it is imperative that the data be both plentiful and accurate. The accuracy of the data used in this study was not rigorously tested. Rigorous testing would include such activities as taking a random sample of the actual Coast Guard accident reports used to create the data base and checking the data from the reports with the coded data on the computer file. Instead, a heuristic type of testing was done. First, the computer data base was checked for duplicates. Two duplicates were found, which were promptly removed. Second, data items were reviewed for obvious errors. An example of an obvious error would be an operator listed as being two years old. Fortunately, this particular type of error did not occur, and in fact very few of the data items were found to be miscoded. The real problem was not erroneous data, but lack of data. In particular, many records contained missing data items. Table 13, below, gives the percentage of records that contained missing data for a randomly-selected group of data items.

Table 13 Percentage of Records Missing Selected Data Items

item	*data*	*percentage*
13	Operator's years of experience	39
14	Operator's hours on duty	39
26	Number of loaded barges in the towboat-barge combination	24
33	High or low water at the time of the casualty	86

Additional problems included inaccuracies and inconsistencies with the measures of severity. Item 61 and Item 69 suffer from inaccuracies. They are inaccurate because they are not actual dollar damage figures for the accident, but are, as explained above, estimates made by Coast Guard investigating officers. Item 64 suffers from inconsistencies. It is inconsistent because it depends on the Coast Guard investigating officer's interpretation of the Federal regulations. He must interpret these regulations to decide whether or not a vessel casualty is reportable.

In terms of reflecting or capturing reality, the number of vessel casualties is probably the better measure of severity. The main reason for this is that it is easier to determine. For example, the Federal regulations specify that there must be at least $1,500 in damages to make a casualty reportable. It is much easier for an investigating officer to judge whether or not a vessel has sustained more than $1,500 in damages, and thus to determine whether or not a vessel casualty is reportable, than it is for him to specify the exact dollar damage figure.

One final point on the data needs to be made. In the statistical tests and statistical models presented in the next chapter, a basic assumption is made regarding the relationship between the accident sample used for the study and the actual population of all towboat accidents. This assumption is that the sample being used accurately reflects, and is essentially the same as, the population of all towboat accidents. Although one of the statistical tests is oriented toward this assumption, the assumption was not formally tested. On the other hand, the sample was checked by reviewing the method used for collecting the sample in Coast Guard Report No. CG-D-80-78 [Ref. 12] and through conversations with people who created

the sample. Since no bias was found, the statistical analysis was conducted under the assumption that the sample was a good representation of the actual population of all towboat accidents.

IV. STATISTICAL ANALYSIS

A. STATISTICAL TESTING

This part of the statistical analysis is a straightforward application of various statistical tests to the data. All of the tests used are fully explained in either Dixon [Ref. 2] or Siegel [Ref. 7]. Each application of a test is organized in similar fashion. The organization consists of five parts: (a) Question; (b) Statistical Test; (c) Test Procedure; (d) Statistical Conclusion; and (e) Answer. Interspersed among the parts are various comments about the test, under the heading "Remarks."

In general, the procedure used in developing this section was to pose a question about towboat accidents and then to use a statistical test to answer the question. Before a statistical test was used, a careful check was made of the assumptions required by the test. If the test was found to be valid and the assumptions of the test were met, the test was applied.

For each statistical test a null hypothesis, or a hypothesis under which the test was conducted, was specified. Also specified was an alternate hypothesis, or a hypothesis to be accepted if the null hypothesis was rejected. In the test applications found in this section, the specification of these hypotheses is essentially a rewording of the original question into two competing answers to that question.

Whether or not the answer to the question specified by the null hypothesis was accepted depended both on the outcome of the test and on the rejection level or significance level chosen. The significance level chosen for the tests applied in this section was one-tenth. This level is commonly

the alpha level, and it is based on the maximum probability of making the error of rejecting the null hypothesis when it should be accepted. The choice of the alpha level was based on both past experience with statistical analysis and on past experience with the investigation of marine accidents.

The idea of the significance level is easy to understand if it is thought of in terms of chances or odds. For example, if a one-tenth level of significance is used in judging a hypothesis, it means that the chances are one out of ten that the hypothesis would be rejected when it should be accepted. In other words, it is 90% certain that the right decision has been made.

The most important aspect of the statistical testing approach of this section is the relationship between the probability level computed by the test and the two competing hypotheses or answers to the question. The relationship is important because the probability level computed by the test determines which answer or hypothesis to accept and which to reject. The decision is made when the test level is compared against a predetermined significance or odds level--one-tenth in the case of the tests below. In more general terms, the statistical testing approach allows us to differentiate mathematically between the two competing hypotheses. Although for most of the tests human intuition would give the correct answer, the statistical testing approach is important because it confirms that answer.

1. *Statistical Test One*

Question: For the accident sample, is there a random order between the days on which accidents occurred and the days on which they did not occur?

Statistical Test: This question was answered through the application of the one-sample runs test. The basic requirements of the test were met. In brief, there is a single sequence of observations, days of the week, which has two categories, days with accidents and days without accidents. Furthermore, this sequence of days exhibits runs in which there are several days without an accident and then several days with accidents.

Test Procedure: The null hypothesis states: The days on which accidents occur and do not occur are in a random order. The alternate hypothesis states: The days on which accidents occur and do not occur are not in a random order. The specification of the alternate hypothesis is non-directional, making this a two-tailed test.

The data was broken down by years. For each year, the number of days with accidents and the number of days without accidents were counted. Also counted for each year was the number of runs of days with accidents and the number of runs of days without accidents.

A run can be illustrated by considering a hypothetical year in which accidents occurred on every day of the year for the first half of the year and in which no accidents occurred on any day for the last half of the year. For the hypothetical year, two runs would be counted, one run consisting of the days with accidents and the other run consisting of the days without

accidents. Of course, for an actual year from the sample, the runs would be much shorter, adding up to many more runs during the year. The counts for the various years are displayed in Table 14, below.

Table 14   Counts for the One-Sample Runs Test

year	*days with accidents*	*days without accidents*	*days in the year*	*number of runs*
1971	45	320	365	85
1972	88	278	366	137
1973	86	279	365	136
1974	91	274	365	133
1975	111	254	365	153
1976	71	295	366	102
Total	492	1,700	2,192	746

The standard procedure for the one-sample runs test calls for computing a test statistic that comes from a normal distribution. A normal test statistic was computed for each category of years, including the category of "all years," or the total accident sample. Also computed was the probability of obtaining a normal statistic equal to or more extreme than the test statistic. The computations are displayed in Table 15, below.

Table 15 Results of the One-Sample Runs Test

year	*normal test statistic*	*probability level*
1971	1.24	.11
1972	.33	.37
1973	.52	.30
1974	-.65	.26
1975	-.31	.38
1976	-2.26	.01
All years	-1.11	.13

Statistical Conclusion: Since this is a two-tailed test, the overall significance level of .1 was split, and .05 was placed in each tail of the normal distribution. The only probability level in the above table that falls below .05 is for the year 1976. For this year, the null hypothesis was rejected. For all other individual years, including the category of "all years," the null hypothesis was accepted.

Answer: The answer to the question is that for each individual year in the accident sample except 1976, the days on which accidents occurred and the days on which they did not occur are in a random order. This statement is also true for the entire accident sample, which is the "all years" category in Table 15.

Remarks: A check was made to see how many accidents occurred per day in the accident sample. It was found that approximately 93% of the days on which accidents occurred had only one accident per day. The remaining days had mostly two accidents per day, with very few having more than two.

2. *Statistical Test Two*

Question: For the sample of accidents occurring during 1975, is the frequency of accidents approximately the same for each day of the week?

Remark: The reason for asking this question is to determine whether accidents occur with a greater frequency on certain days of the week. For example, are accidents more likely to occur on weekends? The year 1975 was chosen over the other years for this question because of the results of Statistical Test One. In that test, the days on which accidents occurred and the days on which accidents did not occur appeared random for 1975.

Statistical Test: The statistical test chosen to answer this question was the chi-square test. The basic requirements of the test were met. In brief, the days of the week are nominal data categories, and frequency counts are being made.

Test Procedure: The null hypothesis states: There is no difference in the expected number of accidents for each day of the week--in other words, the frequencies of accidents for each day of the week are equal. The alternate hypothesis states: The frequencies of accidents for each day of the week are not equal.

A computer program was written to count up the number of accidents on each day of the week for 1975. The counts are displayed in Table 16, below.

Table 16 Accident Counts for Days of the Week in 1975

Mon.	*Tue.*	*Wed.*	*Thu.*	*Fri.*	*Sat.*	*Sun.*
15	18	15	15	25	20	24

If the null hypothesis were true, the expected number of accidents on each day of the week would be equal. An expected number of accidents for each day of the week was computed by summing all the accidents and dividing by seven. This number is 18.86.

The standard procedure for the chi-square test calls for computing a test statistic that comes from a chi-square distribution. A chi-square statistic of 5.88 was computed from the observed and the expected frequencies. The probability of obtaining this statistic, or one more extreme, is .437.

Statistical Conclusion: Using a .1 level of significance in the right-hand tail of the chi-square distribution, the null hypothesis that the accidents occur with equal frequency on each day of the week cannot be rejected.

Remark: A check was made to insure that a large number of accidents on one particular day of the year was not influencing the counts. A count was made of the number of days containing one accident, two accidents, etc. The

results of the count are: 91 days with one accident, 19 days with two accidents, and only one day with three accidents. Thus, a large number of accidents on one day was not influencing the counts.

Answer: For the sample of 1975 accidents, the frequency of accidents for each day of the week is statistically the same.

Remark: Although the accident frequencies are statistically the same, there were nevertheless more accidents on Friday, Saturday, and Sunday than there were on other days of the week. This indicates that there may be some effect on the accident rate of holidays, such as the weekend. It should also be noted that this test does not take into account various factors which may be affecting the accident rate. Examples of such factors would include vessel traffic, operating hours of towboats, etc. If data on these factors were available, it might be possible to design a better test application to determine if the frequency of accidents is different for various days of the week.

3. *Statistical Test Three*

Question: Is the proportion of various types of accidents the same for all the navigable channels in the accident sample?

Remark: Another way of phrasing this question would be: "Is there some relationship between the type of casualty and the river on which the casualty occurred?"

Statistical Test: The statistical test chosen to answer this question was the chi-square test. The basic requirements of the test were met. Briefly reviewing those requirements, the data is nominal in scale, and frequency counts are being made.

Test Procedure: The null hypothesis states: The proportion of various types of accidents is the same for all the navigable channels. The alternate hypothesis states: The proportion of various types of accidents is not the same for all the navigable channels.

The data was placed in a five-by-four contingency table. Each cell of the table represented a frequency count of a river location and a type of accident. All of the river locations in the accident sample were used. However, not all of the various types of accidents were used, because for some types of accidents the frequency count for a particular river location was too low to meet the assumptions of the chi-square test. The frequency counts are displayed in Table 17, below.

Table 17 Counts of Accidents by Type of Casualty and River Location

type	*Mississippi Waterway*	*Ohio Waterway*	*Illinois Waterway*	*Gulf Intracoastal*
Head-on collisions	8	3	6	55
Bridge rammings	63	19	49	48
Lock/dam rammings	22	68	5	16
Ramming of a moored or anchored vessel	22	1	14	31
Groundings	24	18	8	41

The statistical procedure for the chi-square test calls for the computation of a test statistic from the chi-square distribution. For the above table, a chi-square statistic of 216.19 with 12 degrees of freedom was computed. The probability of getting this statistic, or one more extreme, under the null hypothesis is essentially zero.

Statistical Conclusion: Using a level of significance of .1 in the right-hand tail of the chi-square distribution, the null hypothesis was rejected. The alternate hypothesis that the proportion of various types of accidents is different for the navigable channels was accepted.

Answer: The proportion of various types of accidents on the various navigable channels is different. In other words, on the Mississippi a towboat-barge combination is more likely to experience a bridge ramming, while on the Gulf Intracoastal Waterway a towboat-barge combination is more likely to experience a head-on collision.

4. *Statistical Test Four*

Question: Does the severity of accidents in terms of total dollar damage per accident vary by navigable channel for the entire accident sample?

Statistical Test: The test chosen to answer this question was the extension of the median test. The basic requirements of the test were met. In brief, the data is at least ordinal in scale, and the navigable-channel samples used are independent.

Test Procedure: The null hypothesis states: There is no difference in the median dollar damage per accident for the various navigable channels. The alternate hypothesis states: The median dollar damage per accident differs according to the navigable channel on which the accident occurs.

The measure of severity used in this test is the total dollar damage per accident--Item 69, described in the previous chapter. The median total dollar damage per accidents for all accidents in the sample is $5,976. For each navigable channel, the number of accidents that had more dollar damage than the median and the number that had less dollar damage than the median were counted. All cases of identity with the median were considered to be less than the median. The counts are displayed in Table 18, below.

Table 18
Counts of Accidents Above and Below Median Dollar Damage per Accident for the Various Navigable Channels

channel	*# of accidents above median*	*# of accidents below median*
Lower Mississippi	36	16
Upper Mississippi	54	42
Ohio	60	60
Illinois	38	49
Gulf Intracoastal	95	120

Note: Throughout this study, the Lower Mississippi is defined as that stretch between Mile Point 125 and Cairo, Illinois; the Upper Mississippi is defined as lying above Cairo, Illinois.

The standard procedure for the extension of the median test calls for the computation of a test statistic from the chi-square distribution. For the above data, a chi-square test statistic of 13.46 was computed. The chi-square distribution associated with this statistic was determined to have four degrees of freedom. The probability of obtaining this statistic, or one more extreme, was computed as .009.

Statistical Conclusion: Using a .1 level of significance in the right-hand tail of the chi-square distribution, the null hypothesis was rejected. The alternate hypothesis--that the dollar damage per accident differs for the various navigable channels--was accepted.

Remark: It is interesting to note the proportion of accidents above the median dollar damage to those below the median for each of the navigable channels. These proportions are given in Table 19, below.

Table 19
Proportion of Accidents Above to Accidents Below the Median Dollar Damage per Accident for the Various Navigable Channels

channel	*proportion*
Lower Mississippi	2.3
Upper Mississippi	1.3
Ohio	1.0
Illinois	.8
Gulf Intracoastal	.8

From the table, it is obvious that the Lower Mississippi has a much higher proportion of accidents above the median dollar damage than do the other channels.

Answer: The median dollar damage per accident does vary according to the navigable channel on which the accident occurs, and this holds true for the entire accident sample. It is also very likely that, if further tests were conducted on various locations within particular channels, "hot spots," or places with a high number of severe accidents, could be found. It should be noted that this is not necessarily the same as places with a high number of accidents.

Remark: This test does not give the reason why the dollar damage per accident might be different for the various navigable channels. The reason(s) might be larger towboat-barge combinations on certain channels, more dangerous areas on certain navigable channels, or--most likely--some combination of many factors.

5. *Statistical Test Five*

Question: Does the severity of accidents in terms of the number of vessel casualties per accident vary by year for the entire accident sample?

Statistical Test: The test chosen to answer this question was the extension of the median test. The basic requirements of the test were met. In brief, the data is at least ordinal in scale and the yearly samples used are independent.

Test Procedure: The null hypothesis states: There is no difference in the median number of vessel casualties per accident for the various years. The alternative hypothesis states: The median number of vessel casualties per accident differs for the various years.

The measure of severity used in this test, the number of vessel casualties per accident, is the same as Item 64, described in the previous chapter. The median number of vessel casualties for the entire sample is 2.4. For each year, the number of accidents that had more vessel casualties than the median and the number that had fewer than the median were counted. All cases of identity with the median were considered to be less than the median. The counts are displayed in Table 20, below.

Table 20
Counts of Accidents Above and Below Median Number of Vessel Casualties per Accident for All Years

year	*# of accidents above median*	*# of accidents below median*
1971	22	28
1972	46	53
1973	47	55
1974	38	66
1975	59	73
1976	46	40

The standard procedure for the extension of the median test calls for the computation of a test statistic from the chi-square distribution. For the above data, a chi-square test statistic of 5.67 was computed. The chi-square distribution associated with this statistic was determined to have five degrees of freedom. The probability of obtaining this statistic, or one more extreme, was computed to be .34.

Statistical Conclusion: Using a .1 level of significance in the right-hand tail of the chi-square distribution, the null hypothesis was accepted.

Remark: As in Statistical Test Four, it is interesting to note the proportion of accidents above the median number of vessel casualties per accident to the number below that median for each year. These proportions are given in Table 21, below.

Table 21
Proportion of Accidents Above to Accidents Below the Median Number of Vessel Casualties per Accident for All Years

year	*proportion*
1971	.786
1972	.868
1973	.855
1974	.576
1975	.808
1976	1.150

Although the years are not significantly different, 1976 has the largest proportion of accidents above the median number of vessel casualties.

Answer: The median number of vessel casualties per accident does not vary according to year for the accident sample. In other words, it cannot be said for the years in this accident sample that one year has a significantly higher median number of vessel casualties per accident than another year. This suggests that the median number of vessel casualties per accident is uniform over the years tested.

6. *Statistical Test Six*

Question: Does the severity of accidents in terms of dollar damage vary according to the type of casualty?

Statistical Test: The test chosen to answer this question was the extension of the median test. The basic requirements of the test were met. In brief, the data is at least ordinal in scale, and the type of casualty samples used are independent.

Test Procedure: The null hypothesis states: There is no difference in the median dollar damage per accident for the various types of casualties. The alternate hypothesis states: The median dollar damage per accident does differ for the various types of vessel casualties.

The measure of severity used in this test, the total dollar damage of the accident, is the same as Item 69, described in the previous chapter.

The types of casualties used for the test were head-on collisions, bridge rammings, lock and dam rammings, and groundings. A median dollar damage figure of $6,510 was computed for these casualties.

For each type of casualty the number of accidents having more than the median dollar damage and those having less than the median dollar damage were counted. All cases of identity with the median were considered to be less than the median. The counts are displayed in Table 22, below.

Table 22
Counts of Accidents Above and Below Median Dollar Damage per Accident for Selected Casualties

type of casualty	*# of accidents above median*	*# of accidents below median*
Head-on collisions	47	25
Bridge rammings	91	88
Lock and dam rammings	52	59
Groundings	31	60

The standard procedure for the extension of the median test calls for the computation of a test statistic from the chi-square distribution. For the above data, a chi-square test statistic of 16.20 was computed. The chi-square distribution associated with this statistic was determined to have three degrees of freedom. The probability of obtaining this statistic, or one more extreme, was computed to be .001.

Statistical Conclusion: Using a .1 level of significance in the right-hand tail of the chi-square distribution, the null hypothesis was rejected and the alternate hypothesis was accepted.

Remark: As in the previous two tests using the extension of the median test, it is interesting to note the proportion of accidents above the median dollar damage to the number below the median for each type of selected casualty. The proportions are given in Table 23, below.

Table 23
Proportion of Accidents Above to Accidents Below the Median Dollar Damage per Accident for Selected Casualties

casualty	*proportion*
Head-on collisions	1.88
Bridge rammings	1.03
Lock and dam rammings	.881
Groundings	.517

From the above table, it appears that head-on collisions have the greatest proportion of accidents above the median dollar damage for these four types of casualties.

Answer: The median dollar damage per accident varies according to the type of the accident.

7. *Statistical Test Seven*

Question: For bridge-ramming accidents, does the number of accidents which occur during the day equal the number which occur at night?

Statistical Test: The test chosen to answer this question was the binomial test. The basic requirements of the test were met. In brief, the data is in nominal categories, day and night, and frequency counts can be made.

Test Procedure: The null hypothesis states: For bridge rammings, the frequency of accidents during the day equals the frequency of accidents during the night. The alternative hypothesis states: For bridge rammings, the frequency of accidents during the day does not equal the frequency of accidents during the night. Since the alternate hypothesis does not specify the direction of the difference, this is a two-tailed test.

Counts were made of the number of accidents occurring during the day and the number occurring at night. These counts are displayed in Table 24, below.

Table 24 Counts of Day and Night Bridge Rammings

Day	71
Night	104

The binomial distribution for the null hypothesis would have a probability level of .5 and a total of 175 discrete points. In other words, the null hypothesis implies that day and night accidents should both equal approximately 87. The probability of obtaining the above distribution under the null hypothesis is .008.

Statistical Conclusion: Since this is a two-tailed test, the overall significance level of .1 was split, and .05 was placed in each tail of the binomial distribution. Using .05 in the tails, the null hypothesis was rejected and the alternate hypothesis was accepted.

Answer: For bridge-ramming accidents, the frequency of accidents during the night exceeds the frequency of accidents during the day.

Remark: Although the test does not determine why there are a significantly larger number of bridge rammings at night, it does indicate that this is a problem. One possible explanation would be improper or inadequate lighting of bridges at night; another would be loss of depth perception in human vision at night.

8. *Statistical Test Eight*

Question: Is the mean dollar damage in head-on collision accidents different between the year group 1971-1973 and the year group 1974-1976?

Remark: The measure of severity used in this test is the total dollar damage of the accident. This is Item 69, described in the previous chapter. The reason for separating the data into the two year groups was to determine whether the 1973 regulations involving bridge-to-bridge communications and operator licensing had an impact on reducing the severity of accidents involving head-on collisions. Head-on collisions were chosen for this test because these regulations would most likely have had the greatest impact on this type of accident.

Statistical Test: The statistical test chosen to answer this question was the *t*-test between means. The requirements of the test were met. In brief, both samples had normal-type distributions, and the variances were approximately equal. The data also meets the requirement of being at least interval scale. To be conservative, the particular type of *t*-test used assumed that the variances were not equal for the two groups.

Test Procedure: The null hypothesis states: The mean dollar damage of head-on collisions in the year group 1971-1973 is equal to the mean dollar damage of head-on collisions in the year group 1974-1976. The alternate hypothesis states: The mean dollar damage of head-on collisions for the two year groups is different. Since the alternate hypothesis does not specify the direction of difference between the means of the two groups, a two-tailed test was used.

The data was divided into the two year groups and head-on collisions were selected out. For the 1971-1973 group, the mean dollar damage (in thousands) is 10.59, with a variance of 73.28. For the 1974-1976 group, the mean dollar damage (also in thousands) is 11.26, with a variance of 75.86.

The standard procedure for the t-test calls for calculating a test statistic from the t distribution. The statistic computed for this data was .33. The degrees of freedom associated with this statistic is 70. The probability of obtaining this statistic, or one more extreme, from a t distribution with 70 degrees of freedom is .74.

Statistical Conclusion: Since this is a two-tailed test, a .05 significance level in each tail of the t distribution was used. At this level of significance, the null hypothesis was accepted.

Answer: There is not a significant difference in the mean dollar damage of head-on collisions between the year group 1971-1973 and the year group 1974-1976.

9. *Statistical Test Nine*

Question: Are there differences over the years for the various navigational channels in the ratio of the number of barges to the number of towboats making trips on the channels?

Remark: The data for this test came from Reference 9. The confusing terminology in the question, "the ratio of the number of barges to the number of towboats," is being used because this is the exact way the data is presented in Reference 9. In actual fact, one can assume a strong relationship between the towboat-barge ratio and the more understandable terminology, "the mean size of the towboat-barge combination."

Statistical Test: The test chosen to answer this question was the Friedman two-way analysis of variance. The basic requirements of this test were met. Briefly reviewing those requirements, the test requires that the data be ordinal. The data meets this requirement. The test also requires matched or related samples. In this application of the test, the navigable channels were considered the matching criteria. In other words, each channel was considered to be its own control over the various years. This assumption is realistic, because the industries, environmental conditions, and general aspects of a large navigational channel do not change significantly from year to year.

Test Procedure: The null hypothesis states: the ratio of the number of barges to the number of towboats making trips on the navigational channels is approximately the same for all years. The alternative hypothesis states: There is a difference in the ratio of the number of barges to the number of towboats making trips on the navigable channels over the years.

The ratios of the number of barges to the number of towboats making trips on the navigable channels were first determined from the data. These ratios were determined for the years 1972 through 1977 on the Mississippi, Illinois, and Ohio Waterways. The data did not permit the determination of these ratios for the Gulf Intracoastal Waterway. The data in Reference 9 was also split into upbound and downbound trips. An average of the upbound and downbound ratios was used for this test. The results are given in Table 25, below.

Table 25 Ratio of Number of Barges to Number of Towboats

year	Mississippi Waterway	Illinois Waterway	Ohio Waterway	average ratio
1972	2.825	5.771	3.379	3.992
1973	2.867	5.641	3.446	3.985
1974	2.881	5.452	3.317	3.883
1975	2.962	5.779	3.340	4.027
1976	3.256	6.874	3.531	4.554
1977	3.201	6.261	3.564	4.342

The Friedman test requires that the data be ranked. In this application, since the effect of years is being assessed, the data was ranked in each column. Low ratios were given low ranks, and high ratios were given high ranks. The rankings are displayed in Table 26, below.

Table 26 Ranking of Data for the Friedman Test

year	Mississippi Waterway	Illinois Waterway	Ohio Waterway	rank sum
1972	1	3	3	7
1973	2	2	4	8
1974	3	1	1	5
1975	4	4	2	10
1976	6	6	5	17
1977	5	5	6	16

The standard procedure for the Friedman test calls for computing a statistic that has an approximate chi-square distribution. The statistic computed for this data was 11.57. The degrees of freedom associated with this statistic is 5. The probability of obtaining this statistic, or one more extreme, from a chi-square distribution with 5 degrees of freedom is .04.

Statistical Conclusion: At a .1 level of significance, the null hypothesis was rejected and the alternate hypothesis--that the ratio of the number of barges to the number of towboats differs for the various channels--was accepted.

Answer: There are differences over the years 1972 to 1976 in the ratios of the number of barges to the number of towboats making trips on the various navigational channels. In general, the ratios are increasing over this period of time. This is obvious from the average-ratio column of Table 25, or the rank-sum column of Table 26. As explained in an earlier remark, this probably means that the average size of towboat-barge combinations is increasing.

10. *Statistical Test Ten*

Question: For the years 1972-1976, on the Ohio River, is the size of the towboat-barge combinations travelling the waterway equal to the size of the combinations involved in accidents on the waterway?

Remark: For this test, the size of the towboat-barge combination is defined as the number of barges in the combination. Two samples of data were used for this test. One sample came from the U. S. Army Corps of Engineers, and is found in Reference 9. This sample was used to get the number of barges in towboat-barge combinations making trips on the waterway. The other sample was a selection of all the Ohio River accidents from the data base described in the previous chapter. This sample came from the U. S. Coast Guard, and was used to get the number of barges in towboat-barge combinations involved in accidents on the waterway.

Statistical Test: This question can best be answered by means of a statistical test called the Mann-Whitney U test. The basic requirements of this test were met. In brief, this test requires both that the data be of an ordinal scale and that two samples be used. In this application, the data was the number of barges in towboat-barge combinations, which is at least ordinal in scale. The two samples used were the U. S. Army Corps of Engineers data and the U. S. Coast Guard data. These two samples are independent.

Test Procedure: The null hypothesis states: For the Ohio River, the size of the towboat-barge combinations making trips on the waterway is the same as the size of the combinations involved in accidents on the waterway. The alternate hypothesis states: For the Ohio, the size of the towboat-barge combinations involved in accidents on the waterway is greater than the size of the combinations making trips on the waterway.

The application of the Mann-Whitney U test is a tedious process. It involves ranking all of the data elements from both samples from lowest to

highest, and then using the rankings to compute a test statistic from the normal distribution. The rankings will not be given. The normal test statistic was determined to be -2.78; the probability of obtaining this statistic, or one more extreme, is .003.

Statistical Conclusion: Since the alternative hypothesis predicts the direction of difference between the two samples, a one-tailed test was used. Using a .1 level of significance in the tail of the normal distribution, the null hypothesis was rejected and the alternate hypothesis was accepted.

Answer: For the years 1972 through 1976, on the Ohio River, the size of the towboat-barge combinations involved in accidents on the waterway was greater than the size of combinations travelling on the waterway. In other words, on the Ohio, larger towboats were more likely to be involved in accidents.

Remark: The test was only applied to the Ohio River because of the tedious nature of the test. Also, the Ohio River samples appeared good for both the Coast Guard data and the Army Corps of Engineers data.

11. *Summary of the Test Results*

For each statistical test, the results were given in the form of an answer to a specific question. These answers are summarized in Table 27, below.

Table 27 Summary of Test Results

test	*result*
1	Accidents occur on random days in the accident sample.
2	Accidents occur with equal frequency on each day of the week.
3	Different accident types occur with different frequencies for each navigable channel.
4	The median dollar damage per accident is different for each navigable channel.
5	The median number of vessel casualties per accident is the same for each year of the accident sample.
6	The median dollar damage per accident is different for each type of accident.
7	Bridge rammings occur more often during the night.
8	The severity of head-on collisions was the same before and after the implementation of regulations on bridge communications and operator licensing.
9	Towboat-barge combinations travelling on the inland waterways are getting larger.
10	On the Ohio River, larger towboats were more likely to be involved in accidents.

B. STATISTICAL MODELING

This section is concerned with the problem of building statistical models from the accident data. In particular, the problem of predicting dollar damage from an accident, and the problem of predicting vessel casualties from an accident, suggest themselves. Such predictions can be made, given a set of predetermined variables. The technique used is multiple

regression, an explanation of which can be found in Wesolowsky [Ref. 16] or in Mosteller and Tukey [Ref. 4].

Multiple regression requires a dependent variable, or variable to be explained, that is at least on an interval scale of measurement. Three dependent variables are used in the models. The first variable is Item 69, which is the total dollar damage of a towboat accident in terms of 1967 dollars. The second variable is Item 61, which is the total dollar damage to the vessels involved in a towboat accident, again in terms of 1967 dollars. The third variable is Item 64, which is the number of reportable vessel casualties resulting from a towboat accident. All three of these variables were described in detail in the preceding chapter.

Independent variables are required to explain the dependent variable. These variables must also be at least on an interval scale of measurement. The independent variables used in the models were chosen from the data items described in the preceding chapter. The independent variables which are available for the modeling are listed in Table 28, below.

Table 28
List of Independent Variables for Multiple Regression

item	*data*	*item*	*data*
9	Operator's age	28	Total cargo tonnage
13	Operator's years of experience	29	Combination length
14	Operator's hours on duty	30	Combination width
15	Towboat gross tonnage	31	Maximum draft
16	Towboat length	32	Width of the channel at accident location
17	Towboat horsepower	44	Maximum span of bridge struck
21	Towboat draft	47	Wind speed at the time of the casualty
25	Number of barges in the combination	65	Item 17 ÷ Item 25
26	Number of loaded barges in the combination	66	Item 17 ÷ Item 26
27	Number of light barges in the combination	67	Item 32 - Item 30
		68	Item 44 - Item 30

A multiple regression model requires several assumptions. First, the sample drawn must be random. The randomness of this sample was discussed earlier, and, in general, there is no reason to suspect that the sample is biased. The second assumption is that each array of the dependent variable follows a normal distribution for any combination of the independent variables. This assumption can be relaxed if the sample size is sufficiently large. In the regression models used in this analysis, the sample size was

kept sufficiently large. In particular, a ten-to-one ratio was maintained between the number of cases and the number of independent variables, as recommended by Wesolowsky [Ref. 16]. The third and fourth assumptions are that the regression of the dependent variable on the independent variables is linear, and that all the dependent arrays have the same variance. Both of these assumptions were checked and found acceptable through an examination of the residuals.

The *Statistical Package for the Social Sciences* [Ref. 6] was used to do the multiple regression models. In all of the models, a forward (stepwise) inclusion of the independent variables was used. This means that the independent variables were entered into the regression equation from best to worst. The variable which explained the greatest amount of variance in the dependent variable was entered first, the variable that explained the greatest amount of variance in conjunction with the first was entered second, and so on. For each run, appropriate summary statistics and scatterplots of the residuals were requested.

The large number of missing data elements in the sample caused a serious problem in the multiple regression computer runs. When a run was made, it was necessary to eliminate cases or records with missing values in order to have a good sample. If all of the independent variables given in Table 28 were used in the regression model, this elimination of cases with missing values would, on the average, reduce the original sample between eighty and ninety percent. For example, when all bridge-ramming accidents were selected and all the independent variables were entered into the regression, the original sample size of 179 cases was reduced to 22 cases. If the reduction in sample size was too great, as in this case,

repeated partial correlations were run in order to pick out the most important independent variables.

Running repeated partial correlations on the computer was both time-consuming and tedious. The procedure followed was to run a correlation between the dependent variable and all the independent variables. The first variable picked for the regression was the independent variable that had the most significant correlation with the dependent variable. The second variable picked was the variable that had the most significant partial correlation with the dependent variable after the correlation or effects of the first variable were removed. The third variable picked was the variable which had the most significant correlation after the effects of the first and second variables were removed, and so forth. At each stage of the process, the sample size was checked, and the process was finally stopped when the sample size was reduced to the minimum acceptable level. In general, this approach to selecting the best independent variables was found more effective than trying to guess at which ones were best.

It should be noted that attaching some meaning to the regression coefficients of the models is improper. For example, it is incorrect to speculate and give reasons why some coefficients in a model are positive and others are negative. The interrelationships among the variables in the model prevent individual coefficients from being separated out and discussed. A more complete explanation of this point can be found in Reference 4.

Each of the models presented in this part is organized in a similar fashion. There are four basic parts to each model. The parts are entitled: Model Description, Procedure, Statistical Conclusion, and Results. Various

comments are interspersed among these parts under the heading "Remarks." A one-tenth level of significance was again used for statistical testing.

1. *Statistical Model One*

Model Description: This multiple regression model used the entire accident sample. The dependent variable was Item 69, which is the total dollar damage of the accident adjusted for inflation, as explained in the preceding chapter. All of the independent variables were permitted in the regression except Item 66, Item 68, and Item 44. Because of missing values in the independent variables, the original sample size of 574 cases was reduced to 99 cases. The final sample of 99 cases contained data values for each of the independent variables.

Procedure: A forward (stepwise) regression procedure was used. This procedure selected variables to enter the regression based on the contribution of that variable toward explaining the variance of the dependent variable. Since the sample size was reduced to 99 cases, not all of the variables could be permitted to enter. The regression was stopped after five variables had entered. The results of the regression are summarized in Table 29, below.

Table 29
Regression on the Entire Sample
(Dependent Variable is Dollar Damage per Accident)

Variable	*B*	*F*	*D/F*	*probability*
Item 30	4.20	10.24	1/93	.002 *
Item 47	-4.25	2.36	1/93	.13
Item 16	-1.48	1.78	1/93	.19
Item 26	11.91	2.10	1/93	.15
Item 67	-0.049	1.11	1/93	.29

Constant = -82.11

Multiple R = .47
R^2 = .22
F = 5.39
D/F = 5/93
Std. Error = 327.12
Probability = .0002 *

Significance level = .1
Dollar values in thousands

The main result derived from the regression procedure is the column of *B* regression coefficients given in Table 29. If these coefficients can be found significant, or, in other words, different from zero, then a prediction equation can be written for the dependent variable dollar damage per accident. Whether or not the *B* coefficients are significant can be determined through two statistical tests. First, we test all of the coefficients as a group. If the result of this test is significant, then we test each individual coefficient.

Statistical Conclusion: The first null hypothesis to be tested states: All of the *B*-regression coefficients equal zero. The alternate hypothesis states: One or more of the *B*-regression coefficients are not equal to zero. The F statistic associated with this hypothesis is the overall F level of 5.93. Since the probability value associated with this level is significant at the .1 level, the alternate hypothesis that one or more of the *B*-regression coefficients are not equal to zero was accepted.

A null hypothesis can now be stated on each individual *B*-regression coefficient. The null hypothesis for each coefficient states: The value of the *B* coefficient is equal to zero. The alternate hypothesis states: The value of the *B* coefficient is not equal to zero. The F statistic and the probability level associated with each *B*-regression coefficient are given above. Those regression coefficients that are marked with an asterisk are considered to be significant at the .1 level. In other words, the *B*-regression coefficients associated with the asterisked probability levels were considered not equal to zero.

Results: This model was a poor fit to the data. There are many possible reasons for the bad fit. First, the dependent variable (Item 69) is not the actual dollar damage of an accident, but is only an estimated value, as mentioned in the preceding chapter. Also, Item 69 is rather broad and includes all categories of dollar damage--that is, damage to vessels, cargo, and property. Second, the regression is too broad in scope. The problem is that all types of accidents, in all types of geographical locations, are being used as the base. This is too broad or varied to get an accurate regression. Third, information was not available on what the

towboat-barge combination rammed or struck. For example, if a wooden dock was hit, there would probably be far less damage than there would be if the side of a passing tanker was hit.

2. *Statistical Model Two*

Model Description: This multiple-regression model used only bridge-ramming accidents. The dependent variable was Item 61, which is the dollar damage to all the vessels involved in the accident in terms of 1967 dollars. The independent variables placed in the regression were selected by repeated partial correlations, as described earlier. Those variables selected were the total cargo tonnage of the towboat-barge combination (Item 28), the number of loaded barges in the combination (Item 26), the width of the combination (Item 30), and the age of the operator (Item 9). The use of these four variables caused a reduction in the sample size from 179 cases to 96 cases because of missing data values. The final sample of 96 cases contained data values for each of the independent variables.

Procedure: Like Model One, this Model uses a forward (stepwise) regression. The results are summarized in Table 30, below.

Table 30
Regression on Bridge-Ramming Accidents
(Dependent Variable is Vessel Damage per Accident)

Variable	*B*	*F*	*D/F*	*probability*
Item 28	.0029	81.60	1/91	.000 *
Item 26	-3.01	28.51	1/91	.000 *
Item 30	.21	8.63	1/91	.004 *
Item 9	-.37	3.03	1/91	.085 *

Constant = .99

Multiple R = .76
R^2 = .58
F = 32.06
D/F = 4/91
Std. Error = 22.17
Probability = .000 *

Significance level = .1
Dollar values in thousands

Statistical Conclusion: As in Model One, statistical tests were done on the overall regression and on each of the *B*-regression coefficients. The probability level of those coefficients found significant is asterisked. In this Model, all of the coefficients were found significant. This means that each coefficient has some meaning and can be considered different from zero.

Results: Since all of the regression coefficients were found significant, they can be used in stating a predictive equation, as follows:

Item 61 (dollar damage to vessels) =
.99 + (.0029 x Item 28) - (3.01 x Item 26) + (.21 x Item 30) - (.37 x Item 9)

It is important to consider the accuracy of the prediction equation. Statistically, the accuracy is given by the standard error term of 22.17, or $22,170. In order for this term to apply, it must be assumed that the actual values of the dependent variable are normally distributed about the values predicted by the equation. Examination of the residuals showed that this was generally true. Therefore, under this normality assumption, it can be said that approximately 68% of the actual values of dollar damage to vessels falls within ± $22,170 of the predicted values for this accident sample.

There are many factors causing possible inaccuracies in the predictive equation. The most notable of these factors are the reduction in the sample size due to missing values, and the estimated nature of the dollar damage figure. It should also be noted that for unusual accidents, such as those with a very high dollar damage to the vessels, the predictive equation would probably give a more inaccurate prediction.

3. *Statistical Model Three*

Model Description: This multiple-regression model used only lock- and dam-ramming accidents.The dependent variable was Item 64, which is the number of reportable vessel casualties resulting from the accident. The independent variables placed in the regression were selected by repeated partial correlations, as described earlier. The final independent variables selected were the draft of the towboat (Item 21), the experience of the operator (Item 13), and the horsepower of the towboat divided by the number of loaded barges (Item 66). The use of these three independent variables

caused a reduction in the sample size from 111 cases to 38 cases because of missing values. The final sample of 38 cases contained data values for each of the independent variables.

Procedure: As in the previous Models, a forward stepwise regression procedure was used. The results are summarized in Table 31, below.

Table 31
Regression on Lock- and Dam-Ramming Accidents
(Dependent Variable is Number of Vessel Casualties per Accident)

variable	*B*	*F*	*D/F*	*probability*
Item 21	.38	6.67	1/34	.014 *
Item 13	-.037	5.07	1/34	.031 *
Item 66	-.00029	3.98	1/34	.054 *

Constant = -.054

Multiple R = .52
R^2 = .27
F = 4.10
D/F = 3/34
Std. Error = .91
Probability = .014 *

Significance level = .1

Statistical Conclusion: As in the previous Models, statistical tests were done on the overall regression and on each of the *B*-regression coefficients. The probability level of those coefficients found significant was asterisked. In this Model, therefore, all of the coefficients have some meaning and can be considered different from zero.

Results: Since all of the regression coefficients were found significant, a predictive equation can be stated using the coefficients, as follows:

$$\text{Item 64 (reportable vessel casualties)} = -.054 + (.38 \times \text{Item 21}) - (.037 \times \text{Item 13}) - (.00029 \times \text{Item 66})$$

It is important to consider the accuracy of the predictive equation. Statistically, the accuracy is given by the standard error term of .91, or approximately one vessel casualty. In order for this term to apply, it must be assumed that the actual values of the dependent variables are normally distributed about the values predicted by the equation. Examination of the residuals reveals that this was generally true. Therefore, under this normality assumption, it can be said that approximately 68% of the actual values fall within ± one vessel casualty of the predicted values. Additional factors that may decrease accuracy include the small sample size used, the possibility that the sample does not accurately reflect the true population of towboat accidents, and the possibility of inconsistencies in the counting of reportable vessel casualties. It should also be noted that for unusual accidents, such as those with a high number of vessel casualties, the equation may give a more inaccurate prediction.

V. CONCLUSIONS AND RECOMMENDATIONS

This thesis has examined towboat accidents through the application of a few common statistical tests and models. As the work proceeded, certain conclusions and recommendations evolved. These findings loosely fall into two categories: those relating to towboat accidents and those relating to statistical analysis. The first section of this chapter discusses the findings for towboat accidents, and the second section discusses the findings for statistical analysis.

A. CONCLUSIONS AND RECOMMENDATIONS ON TOWBOAT ACCIDENTS

The effects of three major factors on severity of towboat accidents were examined in the statistical tests. These three factors were the location of the accident, the year in which the accident occurred, and the type of accident. It was found that accident location and accident type had a significant effect on accident severity, while the year in which the accident occurred had little effect. This finding has implications for future studies on towboat accidents. Since the accident is dependent on location and type, it would be advantageous to limit future studies to data collected from a particular accident type, such as grounding, which has occurred in similar locations. Since the year in which the accident occurred has little effect on the accident, the data used in a future study could be extended over a considerable time frame without invalidating the results.

Two of the statistical tests were concerned with the number of barges found in towboat-barge combinations, or in other words with the size of the combinations. One test showed that, on the Ohio River, larger towboat-barge combinations were more likely to be involved in accidents. Another test showed that the size of towboat-barge combinations appears to be increasing over the years. The coupling of these two findings will mean future problems for the towboat-barge transportation system. Since the large combinations are getting larger, and large combinations appear to have more accidents, under this assumption we could expect an increase in the accident rate.

From my experience as a Coast Guard marine accident investigator, it seemed that accidents occurred with greater frequency on certain days, and particularly on holidays and weekends. A likely explanation for this increase was the simple addition of recreational boats to the waterways. One of the tests determined whether or not there was a greater frequency of accidents on certain days of the week. Although a statistically significant effect was not found, it was evident from scanning the data that accidents are more likely to occur on the weekends. In particular, it was found that the accident rate was low for Monday through Thursday, and high for Friday through Sunday. Future studies might analyze this in more detail, and prove statistically that there is indeed a weekend or holiday effect.

After reviewing several night bridge collision investigations, it was apparent that transiting bridges at night poses a special hazard for towboat-barge combinations. A statistical test confirmed this by showing that there are significantly more bridge collisions at night than during the day. A standardized solution to this problem is unfeasible because of the unique architectural design of each bridge. For example, a lighting solution

designed for one bridge may be completely inadequate for another bridge. Despite this difficulty, this remains an important area for further Coast Guard research, especially since the Coast Guard carries the general responsibility of properly marking the waterways with navigational aids.

One of the major questions concerning any regulation is whether or not it is effective. In 1973, regulations were implemented on bridge-to-bridge communications and operator licensing. A statistical test was designed to measure the effectiveness of these regulations on the severity of head-on collision accidents. Although it was obvious that both of these regulations were needed, a statistical test showed that the regulations did not reduce the severity of head-on collisions. It is possible, however, that the regulations might have reduced the frequency of head-on collisions. Future efforts may wish to address this issue.

An attempt was made, through the use of multiple regression, to develop equations which predict the severity of towboat accidents. In general, the equations were found to be inaccurate, due to the missing data in the accident sample. This effort did show, however, that it is possible to predict the severity of a towboat accident from certain accident variables. With a large, complete, and accurate data base, future efforts may be able to generate a useful set of predictive equations. Equations of this nature would help towboat companies assess the risks of an accident, help insurance companies estimate claims, and assist government agencies in making better regulatory decisions.

B. CONCLUSIONS AND RECOMMENDATIONS ON STATISTICAL ANALYSIS

An important limitation of the statistical testing approach has been the lack of a complete non-accident data base. Although some non-accident information was used from the Army Corps of Engineers, an accident data base was needed that contained information on accident-free passages on the inland waterways. The addition of a non-accident sample would have added another dimension to the statistical tests. With both an accident and a non-accident sample, statistical tests would be able to discover possible causes of towboat accidents.

A persistent problem with the data base of this study was the question of whether or not the variables in the data base were the best variables. It is entirely possible that data was collected on some meaningless variables, while no data was collected on some of the more important variables. One solution to this problem is the implementation of the Delphi Technique. This technique was used in the Texas Highway Department report [Ref. 8] which identified the variables needed for a computerized data base on highway conditions. Were this technique applied to towboat accidents, the end result would be a list, in order of importance, of all the possible accident variables. This list could then be used to determine on which variables data should be collected.

An important limitation in applying the multiple regression technique was the substantial amount of missing data. Since the accident data base was created from investigative reports, this problem point to some deficiency in the way accident information is collected. One possible solution to this problem is to introduce a quality control system which assigns grades to accident investigations. Using the importance scale of the data

variables given by the Delphi Technique, a numerical importance value can be assigned to each variable. For example, more important variables could be given high numbers, and less important variables could be given low numbers. From the importance number of each variable, a grade can be computed for an accident investigation by simply adding up the importance numbers of those variables on which data was collected. An investigation with a high grade would then have more important accident information than an investigation with a low grade. This scoring system would have an effect similar to the grading system commonly used in schools. More precisely, those investigators who were concerned about the grade on their reports would do a better job of both investigation and of collecting data.

During the course of this effort, several statistical techniques were examined for use on the accident data. One of these techniques was contingency-table analysis. Although this technique was not found suitable for analyzing the towboat accident data, it may prove very useful in the analysis of recreational boating accident data. The technique appears to be particularly suited to boating accidents because the dependent accident variables, such as death or injury, and the independent accident variables, such as manufacturer of the boat and type of boat, are commonly categorical in nature. In particular, this statistical technique could be used to identify high-risk boating situations in which deaths are more likely to occur. A good example of the actual use of this technique can be found in Reference 5.

In general, the statistical analysis approach was found to be an advantageous way to study towboat accidents, its main advantage being that

it is mathematically defensible. In other words, the hypothesis accepted or the answer chosen was based on a series of logical mathematical steps that can be easily verified and defended.

Statistical analysis using a computerized data base is a difficult task, and it cannot be halfheartedly undertaken. Yet, currently, it is done on a part-time basis by many branches at Coast Guard Headquarters. An effective example of a branch exclusively devoted to statistical analysis is the already-functioning Mathematical Analysis Division of the National Highway Traffic and Safety Administration (NHTSA). The Coast Guard's statistical analysis efforts would be both more accurate and more effective if they were done, like those of NHTSA, on a full-time basis by an experienced staff in a designated branch or division.

As a result of this study, four major suggestions for improving data collection and analysis on towboat accidents seem valid: (a) use the Delphi Technique to identify the important variables in towboat accidents; (b) initiate a quality-control system on the data collected from accident investigations; (c) establish a non-accident data base for comparison with the accident data collected; and (d) establish a statistical analysis branch at Coast Guard Headquarters to analyze data and to provide feedback on the quality of data collected.

These suggestions come from the experience of using a large-scale computer on accident data. Since the Coast Guard presently uses computers for accident data storage and retrieval, and will soon be using computers to collect data under the Marine Safety Information System network, these suggestions are even more relevant, not only for towboat accidents but for all types of marine accidents. In particular, the adoption of these four

suggestions will help the Coast Guard control the problem of marine accidents and will generate better decisions on when to regulate and when not to regulate towboats and other vessels using the navigable waters of the United States.

LIST OF REFERENCES

1. The American Waterways Operators, Inc., *Big Load Afloat*, 1973

2. Dixon, W. J., and Massey, F. J. Jr., *Introduction to Statistical Analysis*, third edition, McGraw-Hill, 1969

3. Howe, Charles W., et. al., *Inland Waterway Transportation*, The Johns Hopkins Press, 1969

4. Mosteller, F., and Tukey, John W., *Data Analysis and Regression*, Addison-Wesley, 1977

5. National Highway Traffic and Safety Administration Technical Note DOT HS-802 332, *The Severity of Large Truck Accidents*, by James Hedlund, April, 1977

6. Nie, N. H., et. al., *Statistical Package for the Social Sciences*, McGraw-Hill, 1970

7. Siegel, Sidney, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, 1956

8. Texas Highway Department Report 123-27, *Influence of Cognitive Style in a Methodology for Data Base Design*, by Ronald R. Bush, February, 1975

9. United States Army Corps of Engineers, *Waterborne Commerce of the United States*, 1972-1977

10. United States Coast Guard Comdtinst M16754.1, *Boating Statistics, 1978*, 1 May 1979

11. United States Coast Guard, *Coding Instructions for Commercial Vessel Casualties*, 1976

12. United States Coast Guard Report No. CG-D-80-78, *Human and Physical Factors Affecting Collisions, Rammings, and Groundings on Western Rivers and Gulf Intracoastal Waterways*, by B. Paramore and others, January, 1979

13. United States Coast Guard Comdtinst M16700.2, *Marine Safety Statistical Review*, 1 January 1979

14. United States Coast Guard, Office of Merchant Marine Safety, *Statistics of Casualties*, 1971-1978

15. United States Department of Commerce, *Statistical Abstract of the United States*, GPO, 1978

16. Wesolowsky, George O., *Multiple Regression and Analysis of Variance*, Wiley, 1976

INITIAL DISTRIBUTION LIST

		No. copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Professor D. E. Neil, Code 55Ni Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
4.	Commander (dpl) Thirteenth Coast Guard District 915 Second Avenue Seattle, Washington 98104	2
5.	American Waterways Operators, Inc. 1600 Wilson Boulevard Arlington, Virginia 22209	2
6.	Commandant (G-DST-1) U. S. Coast Guard Headquarters 2100 Second Street S. W. Washington, D.C. 20590	2
7.	Commandant (G-PTE-1) U. S. Coast Guard Headquarters 2100 Second Street S. W. Washington, D.C. 20590	2
8.	Department Chairman, Code 55 Department of Operations Research Naval Postgraduate School Monterey, CA 93940	1

Thesis 188848
G1425 Gamble
c.1 An analysis of
accidents involving
towboat-barge combina-
tions on selected in-
land waterways of the
United States.